Chapter 1

The Pleiades as a benchmark

1.1 Generalities

Description of the current knowledge on the pleiades cluster Age, distance, metallically reddening, number of members, proper motion. Describe what are they.

1.1.1   The distance controversy

1.2   Spatial Distribution

1.3   Velocity Distribution

1.3.1   Radial Velocity

1.4   Luminosity Distribution

1.5   Mass Distribution

1.6   The current dynamical scenario

1.7   The Pleiades DANCe DR2.

This section must contain a detailed description of the DR2 data.

Table 1.1: Pleiades DANCe DR2 properties

1.7.1   Particularities of the Pleiades DANCe DR2

As described in Section **??** the Pleiades DANCe DR2 contains astrometric (stellar positions and proper motions) and photometric ($ugrizYJHK_s$) measurements for 1,972,245 objects.

1.7.2   Selection of variables

Sarro et al. (2014) demonstrated that the most effective variables for the discrimination of members are the proper motions and the $riYJHK_s$ bands. However, they excluded the r band due to its large number of missing values in their training

set. The selection of variables in this work aims at comparing its results with those found by Bouy et al. (2015b) using the $\mu_\alpha, \mu_\delta, J, H, K_s, i - K_s, Y - J$ variables.

The set of variables used in this work are the stellar positions, the proper motions in right ascension and declination, μ_α, μ_δ, and the photometric colours and magnitudes, $i - K_s, Y, J, H, K_s$. However, in order to compare the results with those of Bouy et al. (2015b), the analysis of the spatial distribution of the Pleiades stellar positions is done independently, in Olivares & et al. (2017b).

As described in Olivares & et al. (2017a), the photometry is modelled by parametric series of cubic spline. The parameter of this series is the colour index $i - K_s$ (in the following CI). This colour allows the most one-to-one variate-covariate relation. Figure 1.1 shows the colour-magnitude diagram (CMD) K vs $Y - J$ the second most one-to-one relation.

Figure 1.1: K vs $Y - J$ CMD for the Pleiades candidate members of Bouy et al. (2015b)

.

1.7.3 Data preprocessing

Since both photometry and proper motions carry crucial information for the disentanglement of the cluster population, we restrict the data set to objects with proper motions and at least two observed values in any of our four CMDs: Y, J, H, K_s vs CI. This restriction excludes 22 candidate members of Bouy et al. (2015b), which have only one observed value in the photometry. Furthermore, we restrict the lower limit ($CI = 0.8$) of the colour index to the value of the brightest cluster member. We do not expect to find new bluer members in the bright part of the CMDs. We set the upper limit ($CI = 8$) of the colour index at one magnitude above the colour index of the reddest known cluster member, thus allowing for new discoveries. Due to the sensitivity limits of the DR2 survey in i and K_s bands, objects with $CI > 8$ have K_s magnitudes ≥ 16 mag. These objects are incompatible with the cluster sequence and therefore we discard them a priori as cluster members.

Our current computational constraints and the costly computations associated

to our methodology (described throughout this Sect.), prevent its application to the entire data set. However, since the precision of our methodology, as that of any statistical analysis, increases with the number of independent observations, we find that a size of 10^5 source for our data is a reasonable compromise. Although a smaller data set produces faster results, it also renders a less precise model of the field (in the area around the cluster) and therefore, a more contaminated model of the cluster. For these reasons, we restrict our data set to the 10^5 objects with highest membership probabilities according to Bouy et al. (2015b). Of this resulting data set, the majority ($\approx$98%) are field objects with cluster membership probabilities around zero. Thus, the probability of leaving out a cluster member is negligible. For the remaining of the objects in the Pleiades DANCe DR2, we assign membership probabilities *a posteriori*, once the cluster model is constructed.

Chapter 2

Results

In this Chapter I present the results of applying the methodology explained in Chapter ?? to the Pleiades DANCe data set (Sect. 1.7). However, to characterise the methodology and estimate its precision and accuracy, I first apply it over synthetic data. Using this data, I am able to analyse the performance of the methodology when it is considered only as a classifier. Later, in this Chapetr I give the results of our methodology main objective: the statistical characterisation of the cluster population. In the following Sections I give the details of the spatial, velocity, luminosity and mass distributions. Later, I will describe the physical scenario of the evolution of the mass distribution by comparing the Pleiades mass distribution with other younger and older clusters. Finally, I will end this Chapter with a description of how the Bayesian methodology allowed us to update our previous knowledge of the Pleiades cluster.

2.1 Performance of the classifier

As mentioned earlier, the main objective of our the methodology of the BHM is the statistical characterisation of the NYOC populations. As a by product, it also obtains individual membership probability distributions. Using these last ones, we are able to directly classify objects into cluster and field members, providing that and objective probability threshold has been stablished. The objective of this section is to find this objective threshold by means of synthetic data. As any other measured property, this classification has an uncertainty, thus, the purpose of this

section is also to quantify this uncertainty.

That said, in order to properly characterise our classifier, I test it over synthetic data sets that resemble the real data. An ideal test to our classifier will be to apply it over well known dataset in which tags of cluster and field members were already present. However, if we may have access to these tags, a classifier may not be needed. The Pleiades cluster being *the* most studied cluster in history, is the NYOC with most of these tags. This is the reason for which we decided to benchmark our methodology on it. Nevertheless, the problem of the synthetic data remains since the low-mass end of the cluster still is *terra ignota*. To over come this issue, we decided to create synthetic data sets under the assumption that our cluster and field models resemble the real data. We are aware that these models are far from perfect, but so far is the best we can do.

The assumption that our model correctly models the real data, although enable us to measure the uncertainty of our classifiers, does not give any indication about possible biases in our model. To explore this possibility, we later compare our real data results with those found in the literature. I present this comparison at the end of this section.

Hence, we fit our models, field and cluster, to the real data (10^4) and using the MAP estimates, we created synthetic stars (five samples of 10^4 objects each). To further test the reliability of our classifier, we compare the results it render when applied over data sets with and without missing values. This comparison allows us to quantify the impact that missing values have on our results.

One further consideration. The synthetic analysis requires at least three runs: one on the real data to obtain the MAP, and two on the synthetic one: with and without missing values. However, as explained in Chapter **??**, our methodology is computationally expensive. Therefore, to maintain the computing time within reasonable limits (couple of weeks), we decided to restrict our synthetic data set to only the 10^4 objects with higher membership probability according to Bouy et al. (2015b). I elaborate on the consequences of this decision.

These 10^4 objects are "closer" to the cluster in the sense of membership probability than the remaining $9{\times}10^4$ objects. Therefore, the field probability density rendered by this sample, compared to that of the larger 10^5 sample, has the following properties: it is more concentrated and has larger values near the cluster

region. Since the 10^5 sample is largely dominated by the field population, its density peaks far form the cluster proper motion and photometric sequences. Densities are normalised, thus more of the density mass of this large sample is far from the cluster region.

Given the previous considerations, we assume that results obtained on the smaller 10^4 sample have are more contaminated, and have lower recovery rates than those obtained on the larger and more distant 10^5 sample. On the one hand, the higher contamination rate results from the larger concentration of the field density around the cluster region. On the other hand, the lower recovery rates arises from the higher values of the field probability density. In simple words, when we define the field in a more restricted region around the cluster, both populations become more entangled, thus they become more difficult to separate. Therefore, we assume that the results obtained on the smaller 10^4 sample represent upper and lower limits to the contamination and recovery rates of the larger 10^5 sample, respectively.

Briefly, to create the synthetic data set, the procedure is the following. First, using the methodology of the previous Chapter, I obtain a sample of the posterior distribution of the parameters given the 10^4 real data set. Then, I chose the particle with highest posterior probability as the MAP estimate of the posterior distribution. Using this particle positions I generate five synthetic data sets of 10^4 objects each. Then, I tag these objects according to their parent population: cluster or field. Then, using the synthetically observed values I estimate their uncertainties and missing value patterns (more details below). Finally, I run the model over these five synthetic samples and compare the measured tags with the true ones as function of the probability threshold.

Also, I run the methodology over the synthetic data set without missing values and compare this results with those found on the same data set with missing values. This test, as mentioned before, enable us to quantify the impact of missing values over individual membership probabilities.

As explained in Sect. 1.7, our data set has a high fraction of missing values. Only $\sim 1\%$ has completely observed entries. Furthermore, the missing pattern is not random and depends on the magnitudes and colours of the objects. Therefore, to better reproduce this pattern, for each synthetic datum, we use missing value

template of one of its closer neighbours in the real data; closer in the euclidean sense. Using the missing value template of the nearest neighbour from the real data set results in a biased sample in which objects with complete (non missing) values are underestimated. This is the inevitable consequence of the fact that euclidean distances measured in subspaces resulting from the missing values, are smaller or, at most, equal to those measured in the non-missing value spaces.

Missing values are assigned as follows. Since by definition of our data set there are no missing values in our proper motion data, missing values were assigned only to photometry. We chose the closer neighbours from the available CMDs: $\{K_s, J - K_s\}, \{J, J - H\}, \{K_s, H - K_s\}, \{J, Y - J\}, \{K_s, i - K_s\}$. These CMD are formed with the bands and colours with fewer missing values, in decreasing order. The missing value pattern for individual objects was chosen as follows. First, for each CMD subspace we find the fraction, f_r, of objects from real data without missing values, we call it $C_{or,i}$. Then, we take a random sample from the synthetic data whose fraction, f_s corresponds to f_r. For objects in this sample we assign the missing value pattern of the nearest neighbour from sample $C_{or,i}$. We repeat the procedure for all CMDs. In this way, the synthetic data has fractions of missing and non-missing values similar to those of the real data.

Uncertainties are assigned as follows. We set the proper motions uncertainties to those of the nearest neighbour in the real data. In photometry, however, this scheme renders uncertainties that are biased towards the less precise measurements. This is a consequence of the missing values. Again, the euclidean metric results in the preferential choosing of objects with missing values. These missing values occur mostly at the faint end, where uncertainties are larger. Therefore, the uncertainties are biased towards larger values. To avoid this issue, we fit polynomials (8th degree) to the uncertainties as a function of the magnitudes. Then, we use these polynomials to give uncertainties to the synthetic photometric data.

The performance of our classifier was measured by counting the true positives (TP, cluster members correctly classified), true negatives (TN, field members correctly classified), false positives (FP, field members classified as cluster members) and false negatives (FN, cluster members classified as field members) recoveries as a function of the probability threshold. With them we calculate the true positive rate, contamination rate, accuracy and precision, which are defined as follows. In

order to classify the objects as cluster of field members we summarise their membership probability distribution using the mode. If the mode is grater than the current probability threshold, then the object is classified as cluster member, if not as field.

The true positive rate (TPR) is the ratio of true positives over the sum of true positives plus false negatives. The contamination rate (CR) is the ratio of false positives over the sum of false positives plus true positives. The precision or positive predictive value (PPV) is the ratio of true positives over the sum of true positives plus false positives. Finally, the accuracy (ACC) is the ratio of the sum of true positives plus true negative over the sum of true and false positives and negatives. These are,

$$TPR = \frac{TP}{TP + FN}$$
$$CR = \frac{FP}{FP + TP}$$
$$PPV = \frac{TP}{TP + FP}$$
$$ACC = \frac{TP + TN}{TN + FN + TP + FP}$$

We use the results of the five synthetic data sets to quantify the uncertainties of the previous quantities.

In Fig. **??** I show the TPR and CR, as a function of probability threshold, computed for the cases in which missing values where both present and absent. The missing value case was computed for the five synthetic samples, thus, the lines and the shaded grey regions depict the mean and maximal deviations, of the results found on the five synthetic samples. As can be seen in this Figure, the missing values have a negative impact in our classification. They cause a diminishing of the TPR and an increase in the CR. This negative impact is expected since the observables we are using are highly discriminant in the classification process. Since cluster and field are highly entangled in our 10^4 synthetic samples, when a highly discriminant observable is missing the classification process could be biased or more uncertain.

In spite of the negative impact of missing values, our methodology delivers low

($\lesssim 8\%$) contamination rates and high recovery rates. In Figure **??** we also show , for the sake of comparison, the CR and TPR of Sarro et al. (2014) (reported in their Table 4).

From the previous comparison we observe that the TPR delivered by of our methodology: i) when measured on data without missing values, is similar to that of Sarro et al. (2014) methodology. This is expected since the results of those authors are based on a model constructed only with completely observed objects (i.e. non-missing values). ii) when measured on data with missing values, is $\approx 4\%$ lower than that of Sarro et al. (2014).

This last figure is a small price to pay compared to the dramatical increase in the number of sources used to construct our model (both field and cluster). As mentioned in Section 1.7, in our 10^5 objects data set, only $\sim 1\%$ of objects have completely observed entries. Roughly, this is the fraction of sources used by Sarro et al. (2014) and Bouy et al. (2015b) to construct their models.

On the other hand, the CR of our methodology, above $p = 0.8$ and for both missing-values and non-missing values cases, outperform, by $x\%$, the CR reported of Sarro et al. (2014) methodology, see below.

In simple words, what can be observed from the previous analysis is that our methodology, compared to that of Sarro et al. (2014), renders less contaminated (CR) results ($x\%$) at the price of a smaller recovery rate (TPR) 4%

Nonetheless, we stress the fact that the previous comparison is not straight forward. The following reasons must be born in mind. First, Sarro et al. (2014) infer their cluster model using only non-missing-value objects, later they apply that model to objects with and without missing values. Second, their synthetic data set and ours are essentially different. They are constructed with different generative models, different number of elements, and different missing value patterns.

Now, I describe the procedure to set an objective probability threshold. This probability threshold, although is not needed to obtain the distribution of the cluster population, is needed however to objectively classify an object. Since real data contain missing values, there is no need establish this threshold for the non-missing values case. There are different approaches to establish this probability threshold, however I only use the approach of maximum accuracy (ACC).

Figure **??** shows the ACC and the PPV of our classifier when applied on syn-

thetic data with missing values. The lines and the grey regions depict, respectively, the mean and the maximum deviations of the five synthetic data set results. We use these last ones as the uncertainty of our estimates. The mean of the highest accuracy, ACC=$96.5 \pm 0.1\%$, happens at probability threshold $p_t = 0.84$. At this threshold the CR is $4.3 \pm 0.2\%$, the TPR is $90.0 \pm 0.05\%$, the ACC is , and the PPV is $95.6 \pm 0.2\%$.

Figure 2.1: Left: The TPR (solid line) and CR (dashed line) of our methodology when applied on synthetic data sets with and without missing values (red and blue lines, respectively). In black dots we show the TPR and CR reported by Sarro et al. (2014) for their non-missing values model. Right: Accuracy and precision as a function of probability threshold for our classifier when applied on synthetic data with missing values. The higher accuracy is obtained at $p_t = 0.84$ (red dot). In both panels, the grey areas show the maximum deviations from the mean of the results of the five missing-values synthetic data sets.

We investigate further on the impact of missing values, particularly to analyse any possible biases introduced by them. It was done by comparing the mode of the individual membership probability distributions found after fitting the model to data sets with and without missing values. The missing values data set is the result of adding a mask of missing values (as previously described) to the completely observed data set. It means that this data set has identical observed values, except for those masked as missing ones.

In Fig. **??** we show the mode of the membership probabilities of objects with missing values (vertical axis) against those of non-missing values (horizontal axis). As can be seen in this Fig., the missing values impact our results by spreading the membership probabilities. Ideally, we would like to recover membership probabilities following the line of slope one, as in the case of completely observed values (red squares). The most striking deviations come from objects lacking the CI (enclosed in black). Our methodology uses the *true* CI to prescribe the *true* photometry, and the observed CI to constrain the marginalisation integral of the *true* CI. Thus, it is expected that a missing CI will produce a probability spread. These missing CI objects show two different behaviours. In one case, there are sources with membership probabilities $p_{complete} \approx 0$ which have overestimated probabilities in the

incomplete case (vertical axis). In the other case, the sources in the combed area below the line of unit slope have underestimated probabilities in the incomplete case. While the first case contributes to the CR the second one diminishes the TPR. The first case reaches the maximum difference at $p \approx 0$ (difference between red and blue dashed lines in Fig. **??**), thus its impact in our results is marginal. Furthermore, at our objective probability threshold $p_t = 0.84$, these objects have a small impact in our results, they represent only 1.8% of the contamination rate. The box region in Fig. **??** shows them. The second case, however, typify the unavoidable loss of members due to the missing values, 4% at $p_t = 0.84$, given our model and the observables it uses.

The bias introduced by missing values can be estimated using the root-mean-square (rms) of the difference between membership probabilities recovered with and without missing values. The total rms is 0.12. On the one hand, objects with completely observed values (red squares) in both data sets have a rms of only 0.02. On the other hand, objects with missing values have a rms of X. While the rms of objects lacking the CI have and rms of X. The previous effects show an overall agreement between results on data sets with and without the missing values, nonetheless, care must be taken when dealing individually with objects lacking the CI.

As mentioned before, our methodology aims at the statistical distributions of the cluster population. It works by ensuring that each object contributes to the cluster distributions proportionally to its cluster membership probability. In this sense our results are free of any possible bias introduced by hard cuts in the membership probability. Nevertheless, contamination is still present and must be quantified. To quantify it, I compute the expected value of the CR. It is $\langle CR \rangle = 5.8 \pm 0.2\%$. In this expected value, each CR contributes proportionally to the probability threshold at which it is measured.

In machine learning and data mining is sometimes useful to analyse the performance of a binary classifier by the receiver operating characteristic curve, the ROC curve. This plot is a visual diagnostic of the ability of a classifier to to its job. The ROC curve plots the TPR as a function of the FPR for all possible values of the probability threshold. A perfect classifier would be that in which the TPR=1 and the FPR=0 for some probability threshold. On the other hand, a random clas-

Figure 2.2: Comparison between the cluster membership probabilities recovered from the synthetic data with missing values (Incomplete) and without them (Complete). The colour and shape indicate the amount of missing values. The symbols enclosed in black indicate a missing CI. The top left box contains objects considered as contaminants due to missing values, at the probability threshold $p_t = 0.84$.

sifier would be that on which the TPR=FPR at all probability thresholds. Such classifier has a line of slope one as its ROC curve. Furthermore, the quantitative diagnostic for a binary classifier is the area under the ROC curve (AUC). As its name indicate, the AUC is the integral of the ROC curve. Thus a random classifier has a AUC of one half, while a perfect classifier as AUC=1. In Fig. **??**, I show the ROC curve for our classifier when applied over synthetic data with missing values. It is the ROC of one of the five synthetic realisations described throughout this section. As can bee seen from this Fig. our classifier does an excellent job, with an AUC=0.99.

Figure 2.3: ROC curve of our classifier when applied over one of the synthetic data sets described earlier. As can be seen the AUC=0.99 diagnose it as an excellent classifier.

2.1.1 Comparison with the literature

In Chap. 1 I mentioned that the most important works on the Pleiades members are those of Stauffer et al. (2007); Lodieu et al. (2012); Sarro et al. (2014); Bouy et al. (2015b). In this section I will compare the list of candidate I members found with the most recent study of the Pleiades members, that of Bouy et al. (2015b). Later, for the sake of completeness, I will compare my list of candidate members with the compilation made by Stauffer et al. (2007).

2.1.1.1 Candidate members from Bouy et al. (2015b)

The methodology of this work, although essentially different from that of Bouy et al. (2015b), the fact that both share the data set and use the same observables, allows a direct comparison between them. When using their objective probability thresholds, as shown by Fig. **??**, both methodologies agree on the outstanding 99.6% of the classified objects. Concerning the candidate members, they also agree on $\approx 90\%$ of them. Nevertheless, the discrepancies are worthy of discussion.

The rejected candidates of Bouy et al. (2015b) (lower right box of Fig. **??**) amount to 12% of their total number of candidate members. This value is 4.7% higher than the contamination rate reported by Sarro et al. (2014): $7.3 \pm 1.4\%$. I will not reject a priori these objects as candidate members because, as mentioned in Sect. **??**, our methodology recovers 4% less members when compared to that of Sarro et al. (2014)(see Fig. **??**). The majority, 85%, of these objects have missing values, which indicates that these objects require future follow up. Even the 37 completely observed and rejected candidates can not be discarded as true members due to the fact that our methodology losses 6% of the true cluster members (see Fig. **??**).

On the other hand, our new candidates (upper left box of Fig. **??**) amount to 10% of Bouy et al. (2015b) candidates. This figure is higher than the $4.3 \pm 0.2\%$ CR reported on the previous Section. Thus, it indicates that up to $\sim 6\%$ of these objects can be true cluster members. From these objects, one half have completely observed values, thus 3% of our new candidates are completely observed objects. This las figure agrees perfectly well with the 3% of non recovered members reported by Sarro et al. (2014).

In Figs. **??** and **??** I show the proper motions and K_s vs $i-K_s$ CMD projection spaces of our new candidates and the rejected ones of Bouy et al. (2015b). In the following I elaborate on their properties.

The new candidate members have proper motions uncertainties whose median, $\overline{\mu_{\alpha,\delta}} = \{1.33, 1.33\}\ mas/yr$, is two times larger than those of the candidate members in common with Bouy et al. (2015b), median $\overline{\mu_{\alpha,\delta}} = \{0.65, 0.65\}\ mas/yr$. Also, as shown by Fig. **??**, the majority of the new candidate members, 148, have probabilities lower than 0.95, are located in a halo around the locus of the cluster proper motions, and on top of the cluster sequence in the K_s vs $i - K_s$ CMD. On

the contrary, the new candidates with probabilities higher than 0.95, which are 39, lay in the centre of the cluster proper motions and fall above the cluster sequence in the K_s vs $i - K_s$ CMD. Thus, I hypothesise that: i) Objects whose photometry is compatible with the cluster sequence but are in the proper motions halo, have higher membership probabilities in our methodology due to the increased flexibility of the cluster proper motions model: it now has four gaussians instead of the two of Bouy et al. (2015b). And ii) objects near the centre of the cluster proper motions but located above the cluster photometric sequence sequence, are multiple systems (probably triple systems which can amount to 4% of the population Duquennoy & Mayor 1991) with an increased membership probability due to our more flexible photometric model of the cluster and equal-mass binaries sequences.

The rejected candidates of Bouy et al. (2015b), as it is shown in Figs. **??** and **??**, have proper motions uncertainties with median $\overline{\mu_{\alpha,\delta}} = \{3.15, 3.19\}\ \ mas/yr$. This value is more than four times larger than that of the candidates in common with our members. Also, these objects are distributed along the cluster sequence. Among these objects, those with a relatively high membership probability occur mostly at the middle of the cluster sequence (green squares of Fig. **??**) while those with lower membership probabilities occur at the bright and faint ends (blue and red triangles of Fig. **??**, respectively). These last regions coincide with those where the missing values happen the most. We stress the fact that Sarro et al. (2014) and later Bouy et al. (2015b) construct their models using only completely observed objects (i.e. those without missing values). For their field models both authors use a sample of $\approx 20,000$ objects. Proceeding in that way, as explained in Sect. **??**, underestimates the field density, particularly in the regions where the missing values are more frequent. Underestimating the field likelihood increases in the cluster field likelihood ratio, therefore it increases also the cluster membership probabilities. Furthermore, the proper motions uncertainties of objects at the bright middle and fain ends, have medians of $\overline{\mu_{\alpha,\delta}} = \{4.0, 4.2\}\ \ mas/yr$, $\overline{\mu_{\alpha,\delta}} = \{2.4, 2.4\}\ \ mas/yr$ and $\overline{\mu_{\alpha,\delta}} = \{3.4, 3.4\}\ \ mas/yr$, respectively. These figures are approximately 6, 4 and 5 times larger, respectively, than those of the candidates in common. These large uncertainties produce a proportional spread of the likelihood distribution thus further reducing the membership probability.

I consider that both the large proper motion uncertainties and field likelihoods

are responsible for the diminished membership probabilities of Bouy et al. (2015b) rejected candidates. However, these rejected candidate members cannot be discarded as potential members. Indeed, at the probability threshold of maximum accuracy, $p_t = 0.84$, the TPR is just $90.0 \pm 0.05\%$. It means that there are still 10% of true members within the rejected candidates. To rule out the possibility that these objects are indeed members we need lower proper motion uncertainties and fewer missing values. Future steps will be taken to try to solve this issue.

Summarising, the discrepancies between our individual membership probabilities and those reported by Bouy et al. (2015b) arise from subtle but important differences. The first of them is the more formal treatment of missing values in our methodology and its inclusion in the field model. Taking into account the missing values has two main consequences. The first of them is that, the new photometric model of the field diminished membership probabilities, particularly in the regions where missing values happen the most. Second, the use of missing values in the construction of the cluster model allow us to include the information of good candidate members that were otherwise discarded a priori. The second difference is the higher flexibility of our cluster model, it allows us to increase the membership probability of the previously discarded candidates. Furthermore, as shown by the red squares in the upper left corner of Fig. **??**, the higher flexibility of our cluster model allow us to include as new candidate members previously rejected objects with complete (non-missing) values.

Figure 2.4: Recovered membership probabilities compared to those of Bouy et al. (2015b). Lines show the 0.75 and $p_t = 0.84$ probability thresholds used in both works. The numbers indicate the new candidate members (top left), rejected candidate members (bottom right), and common candidate members (top right).

Figure 2.5: Proper motion (left) and K_s vs. $i - K_s$ CMD (right) showing the new candidate members found in this work. Captions as in Fig. **??**.

Figure 2.6: Proper motion (left) and K_s vs. $i - K_s$ CMD (right) showing the rejected candidate members of Bouy et al. (2015b). Captions as in Fig. **??**.

Figure 2.7: Proper motion (left) and K_s vs. $i - K_s$ CMD (right) showing the rejected candidate members of Bouy et al. (2015b). The colours and shapes are a proxy for their K_s magnitude.

2.1.1.2 Candidate members from Stauffer et al. (2007)

.

Stauffer et al. (2007) published two list of candidate members. The first one contains 1417 objects compiled from the literature (their Table 2). These objects were classified as candidate members by several authors along modern astronomy. As Stauffer et al. (2007) mention, this list is inhomogeneous, incomplete and certainly includes non members. Their second list contains 55 candidate members (their Table 5) found using infrared photometry, and proper motions.

Cross matching these two lists with the candidate members ofBouy et al. (2015b), shows that only 1288 of the 1417 objects on Stauffer et al. (2007) Table 2 were classified by Bouy et al. (2015b) as candidate members. From these, 1132 objects come from the DANce survey and the rest from Tycho+DANCe, see Appendix B of Bouy et al. (2015b). Also, only 34 of the 55 new candidate members of Stauffer et al. (2007) were classified by Bouy et al. (2015b) as candidate members. In Figures **??** - **??** I show the candidate members of Stauffer et al. (2007), and Bouy et al. (2015b). As mentioned by Stauffer et al. (2007), many of the candidate members in their Table 2 are most probably non members. These objects lay far from the locus of the rest of the candidates. Similarly, the majority of the rejected candidates of Stauffer et al. (2007), 15 of those in their Table 5, lay far from the cluster proper motion and photometric loci.

Concerning our list of candidate members, after cross matching the lists, we recover 1139 of the candidate members from Table 2 of Stauffer et al. (2007). Seven more than Bouy et al. (2015b). Also, we recover the same 34 candidates that Bouy
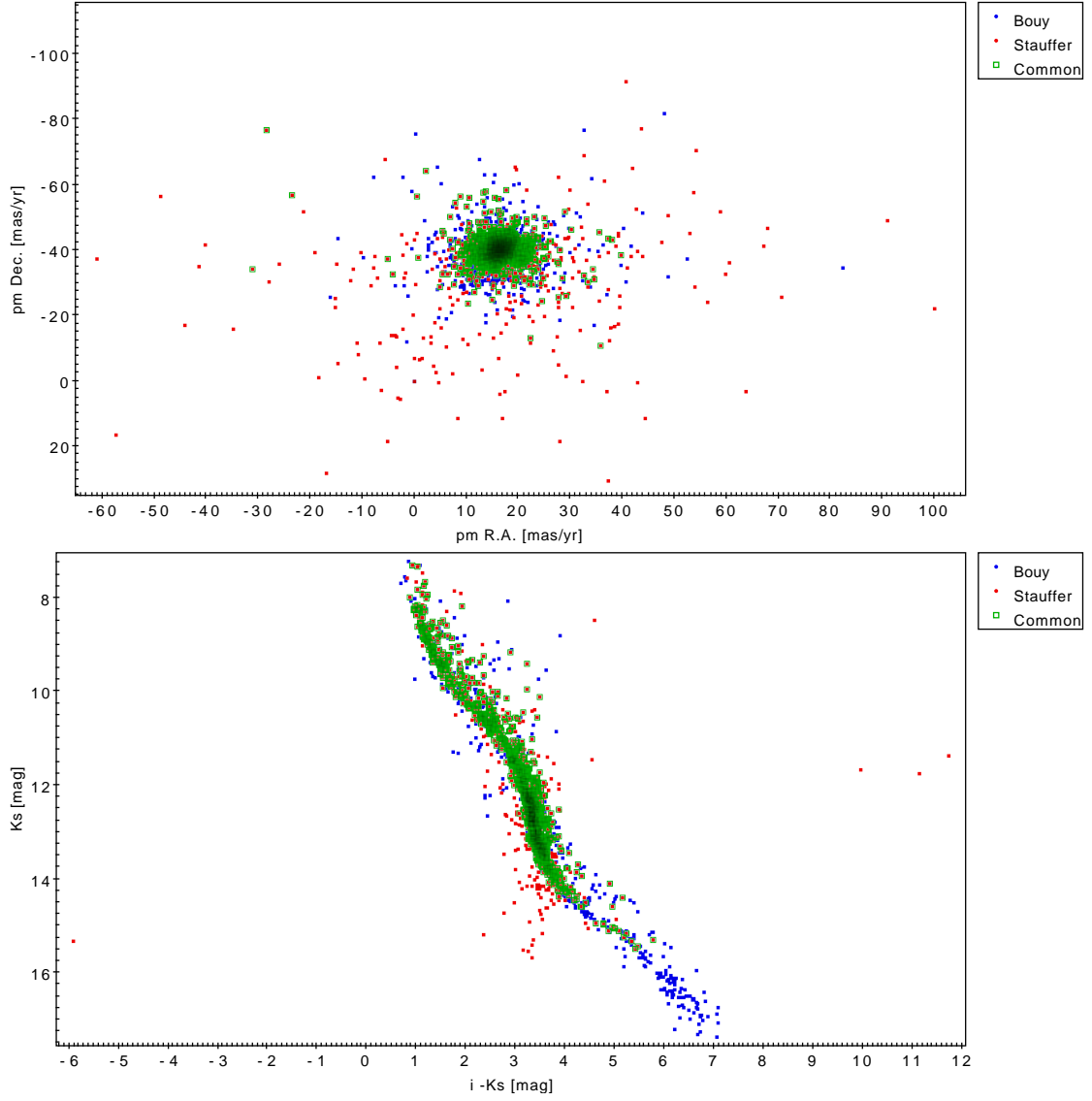
Figure 2.8: Comparison between proper motions (top) and K vs $i - K$ CMD of the candidate members compiled by Stauffer et al. (2007) (red) and those of Bouy et al. (2015b)(blue), the common candidates are shown within green squares.

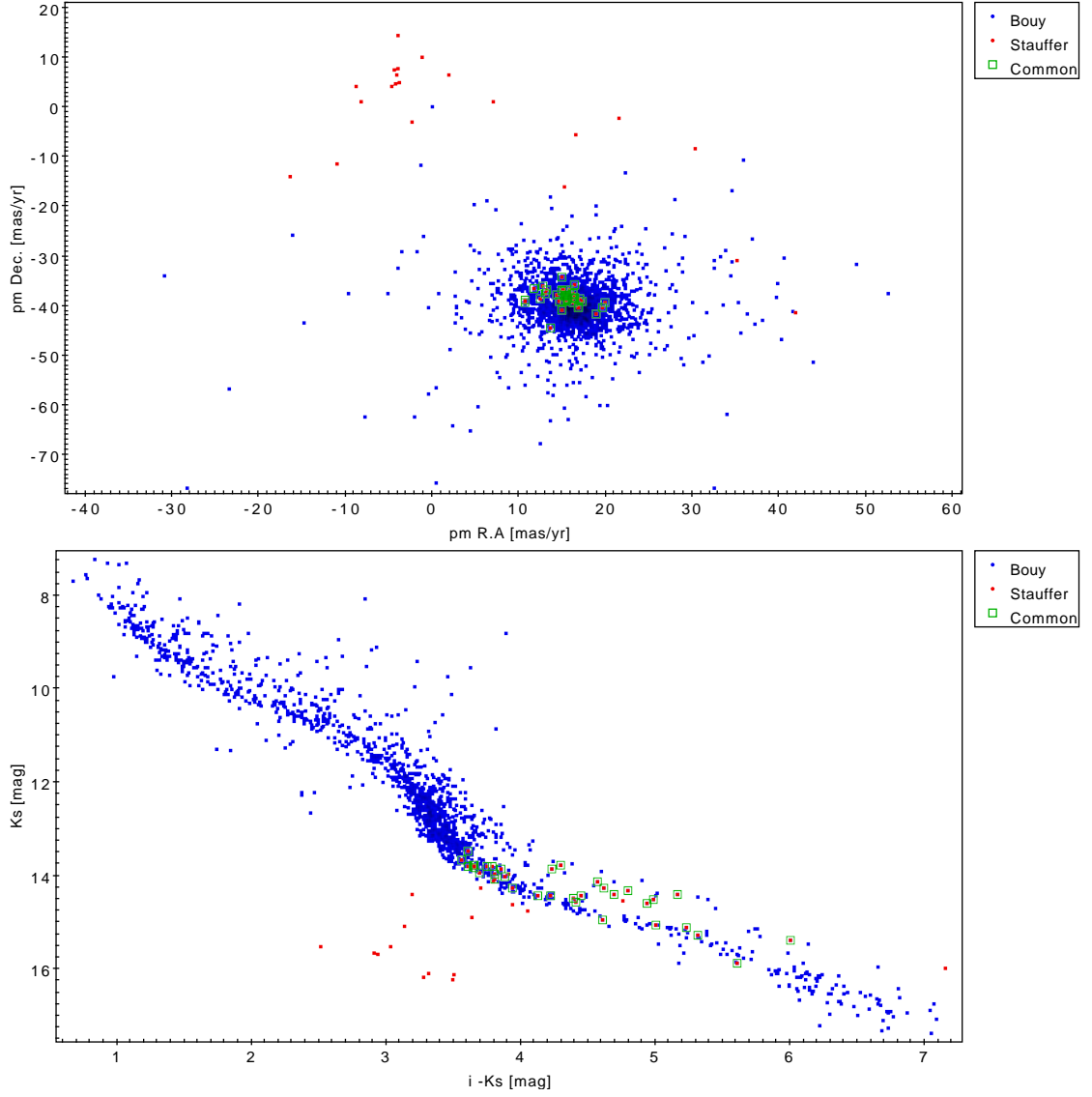Figure 2.9: Comparison between proper motions (top) and K vs $i - K$ CMD of the new candidate members of Stauffer et al. (2007) (red) and those of Bouy et al. (2015b)(blue), the common candidates are shown within green squares.

et al. (2015b) recovered from the new candidates of Stauffer et al. (2007) (from their Table5).

I have already mentioned that the Table 2 of Stauffer et al. (2007) is an exhaustive compilation of Pleiades members. This list contains objects that authors from the literature classified as Pleiades members even when the membership probability was as low as 0.1 (Stauffer et al. 2007). Therefore, I will not analyse in detail the ~ 300 rejected objects. It suffices to mention that we recover an slightly larger (seven) number of candidates than Bouy et al. (2015b).

On the other hand, the list of new candidate members of Stauffer et al. (2007) deserves further attention and a more detailed comparison. From the 21 rejected objects, 14 of them lay below the cluster photometric sequence and far from the proper motion locus. The remaining seven have membership probabilities near our probability threshold. XXXXX CHECK THIS XXXX

2.2 The statistical distributions of the Pleiades cluster.

Once the objective probability threshold has been established and the results of the classification analysed and compared with the literature. Now I present the results of the statistical distributions that describe the cluster population. These distributions result directly or indirectly from the posterior distribution of the parameters in our model. Here I summarise these parametric posterior distributions, together with some of their correlations. Also, I present and analysis on the way these posterior distribution help us to update our prior knowledge.

In Table **??** I summarise the posterior distribution of the parameters in our model. I use as statistic and uncertainty the mode and the 16th and 84th percentiles, respectively. The parameter names correspond to those given in Sect. **??**.

2.2.1 Updating the previous knowledge

As mentioned by Gelman (2006), the posterior distribution must be inspected to update our previous knowledge. To inspect these posterior distribution, I use the

bare values and the correlations among them. The values indicate, for example, that the number of GMM modelling the proper motions of the single stars is overestimated. The fraction and variance of the last gaussian went both to near zero values. A better model would be that in which no computing power will be lost in inferring negligible parameters.

Another simple inspection tells us that our prior for the cluster field fraction is narrow and its mode lay far from that of the posterior distribution. Although this is not wrong since the prior clearly allow this value, this prior must be made wider. Clearly our prior knowledge underestimated the expected number of cluster members.

Another useful tool to inspect these posterior distributions is the correlation they show with themselves. In Figure **??** I show the correlation matrix of all the parameters in our model. This Fig. shows elements that could be considered to improve our model. For example, the large correlation shown by the high order coefficients of the spline series may indicate that we can save some of these parameters.

Figure 2.10: Correlation matrix of parameters in our model. The colour code indicates the value of the correlation coefficient.

In the following Sections, I use the distributions of the parameters in our model to derive the statistical distributions that describe the cluster population. Some of these distributions have as parameters those inferred in the model, however, other, like the mass distribution require more elaborated derivations.

2.3   Velocity distribution

I must work on it. Here I must transform the bivariate proper motions distributions into a univariate velocity distributions using the distance and assuming spherical symmetry. This distribution will give us some hints to improve the proper motions model. It will be interesting to fit a model, maxwellian for example and see how well it fits. If more models are available we can try to find the best one in terms of Bayesian model selection.

2.4   Spatial distribution

To include

2.5   Luminosity distribution

This Section describes the process by which the apparent J, H, and K_s magnitude distributions are obtained and then transformed into the absolute magnitude distributions. Later, I present these distributions and compare them with those found by Bouy et al. (2015b). Also I will compare them with the ones resulted from only the high membership probability objects.

2.5.1   Derivation of the magnitude distributions

To derive the J, H, K_s magnitude distributions I start with the colour index, CI, distribution. This las one, is described by an univariate GMM whose parameters are inferred in the model. Since we also model the EMB, their fraction and photometric sequence are taken into account.

In the following, I exemplify the process of derivation on the K_s band. Similar transformations apply to the rest of the bands. To obtain the distribution of K_s for the cluster objects, I take the colour index CI as a nuisance parameter, later I marginalise it. Thus,

$$p(K_s|\boldsymbol{\theta}_c) = \int p(K_s, CI|\boldsymbol{\theta}_c) \cdot dCI = \int p(K_s|CI, \boldsymbol{\theta}_c) \cdot p(CI|\boldsymbol{\theta}_c) \cdot dCI.$$

The term $p(K_s|CI, \boldsymbol{\theta}_c)$ corresponds to the GMM modelling the distribution of CI (Eq. ??), while $p(K_s|CI, \boldsymbol{\theta}_c)$ is the probability of K_s given the CI, and the cluster parameters $\boldsymbol{\theta}_c$. The EMB are included with an amplitude equal to their fraction, $(1 - \pi_{CB})$. Thus,

$$p(K_s|\boldsymbol{\theta}_c) = \int \left[\pi_{CB} \cdot p_{Cs}(K_s|CI, \boldsymbol{\theta}_c) + (1 - \pi_{CB}) \cdot p_{Bs}(K_s|CI, \boldsymbol{\theta}_c) \right]$$
$$\cdot \, p_{CI}(CI|\boldsymbol{\theta}_c) \cdot dCI.$$
$$= \pi_{CB} \int p_{Cs}(K_s|CI, \boldsymbol{\theta}_c) \cdot p_{CI}(CI|\boldsymbol{\theta}_c) dCI$$
$$+ (1 - \pi_{CB}) \int p_{Bs}(K_s|CI, \boldsymbol{\theta}_c) \cdot p_{CI}(CI|\boldsymbol{\theta}_c) \cdot dCI.$$

$$(2.1)$$

In this equation, Cs and Bs stand for cluster and EMB sequences, respectively. The terms inside the integrals correspond to Equations **??** and **??**. However, since here I want to derive only the distribution of K_s, I marginalise the rest of the bands. Also, the integration limits must change to those of the truncated colour distribution ($CI_{min} = 0.8, CI_{max} = 8$). Hence,

$$p(K_s|\boldsymbol{\theta}_c) = \pi_{CB} \int_{CI_{min}}^{CI_{max}} \left[\left[\sum_{i=1}^{5} \pi_{CI,i} \cdot \mathcal{N}_t(CI|\mu_{CI,i}, \sigma_{CI,i}) \right] \right.$$
$$\left. \cdot \int_{\tilde{Y},\tilde{J},\tilde{H}} \mathcal{N}(\{CI, \tilde{Y}, \tilde{J}, \tilde{H}, K_s\}|\boldsymbol{S}(CI, \boldsymbol{\beta}), \Sigma_{clus}) \, d\tilde{Y} \, d\tilde{J} \, d\tilde{H} \right] \cdot dCI$$
$$+ (1 - \pi_{CB}) \int_{CI_{min}}^{CI_{max}} \left[\left[\sum_{i=1}^{5} \pi_{CI,i} \cdot \mathcal{N}_t(CI|\mu_{CI,i}, \sigma_{CI,i}) \right] \right.$$
$$\left. \cdot \int_{\tilde{Y},\tilde{J},\tilde{H}} \mathcal{N}(\{CI, \tilde{Y}, \tilde{J}, \tilde{H}, K_s\}|T_{Bs}(\boldsymbol{S}(CI, \boldsymbol{\beta})), \Sigma_{clus}) \, d\tilde{Y} \, d\tilde{J} \, d\tilde{H} \right] \cdot dCI.$$

The derivations of the J and H magnitude distributions are similar to the procedure described for K_s. It is worth of notice that, the derivation process takes into account the EMB and the systems (binaries or multiples) which could have different mass ratios. Therefore, we call them the system magnitude distributions.

The previous magnitude distributions, together with the parallax and extinction of the cluster, allows us to obtain the system luminosity distributions. Here, I assume that the parallax is normally distributed with mean, 7.44 mas, and standard deviation 0.42 mas (Galli et al. 2017). This parallax distribution is then

convolved with the magnitude distributions to obtain the absolute magnitude distributions. Finally, I deredden them employing the canonical value of extinction: $A_v = 0.12$ (Guthrie 1987). This last values was transformed transform into the J, H, K_s extinctions using the extinction law of Cardelli et al. (1989).

Our methodology prescribes the *true* photometric quantities based on the *true* colour index CI. Therefore, the completeness limits of this CI dictate those of the photometric bands. The upper completeness limits that Bouy et al. (2015b) estimate for i and K_s are $i \approx 23$ mag and $K_s \approx 18$ mag (see their appendix A). As these authors mention, due to the heterogeneous origins of the DANCe DR2 survey, the completeness is not homogeneous over its entire area. To overcome this issue, they identified a region, the inner three degrees of the cluster, with homogeneous spatial and depth coverage. Then, they restricted their analysis to this region.

Restricting the sample means that good candidate members are not taken into account. Furthermore, if any dynamical process has been set on the cluster such that the mass distribution of its members is not uniform in the space, then restricting the sample could introduce a bias. One of such dynamical process is mass segregation, which, as suggested by Adams et al. (2001) may have happen in the Pleiades.

Instead of restricting the sample, I assume that the UKIDSS survey, which is the most profound of the DANCe catalogue, provides the homogeneous spatial coverage at faint magnitudes, thus providing the lower limit to the completeness. Then, for the bright end of the survey, I quote conservative completeness limits. Figure **??** shows the K_s and i density for all sources in the Pleiades DANCe DR2. The upper completeness limits correspond to the point with maximum density, $i = 21.4$ mag, $K_s = 18.1$ mag. The density at bright magnitudes shows a sharp decline, probably due to saturation. Therefore, I choose $i = 13.2$ mag and $K_s = 11.0$ mag as the lower completeness limits.

The CI completeness interval is then defined as that of all the points, along the cluster sequence in the K_s vs. $i - K_s$ CMD, for which i and K_s are bounded by their upper and lower completeness limits, respectively. This results in a completeness interval for $i - K_s$ of $2.7 < i - K_s < 5.6$ mag. With this completeness interval and the cluster sequence (the splines), we derive the completeness intervals for the

J, H, K_s. Finally, these intervals were transformed to absolute magnitudes and deredden.

Figure 2.11: Density of all DANCe DR2 sources in K_s vs i magnitudes. Lines show our completeness limits, $13.2 < i < 21.4$ mag and $11 < K_s < 18.1$ mag. The grey area is considered incomplete.

.

In Fig. **??**, I plot the J, H, K_s luminosity distributions together with their completeness limits, hereafter I call these distributions the continuous BHM. For the sake of comparison I also show the following luminosity distributions. The luminosity distributions of objects whose mode of membership probability is grater than our probability threshold $p_t = 0.84$, I call these distributions the discrete BHM. Also, I plot the luminosity distribution resulting from the candidate members of Bouy et al. (2015b), I call it discrete Bouy. Since the discrete luminosity distributions, both Bouy and BHM, relay on the individual object magnitudes, and many of these objects have missing values, I impute them using the nearest euclidean neighbour.

The difference between the continuous BHM and discrete BHM distributions comes essentially from the objects used to obtain them. The continuous one uses all objects proportionally to their cluster membership probability while the discrete BHM uses only the high probability candidate members ($p > p_t$). Since the discrete BHM is not a random sample of the continuous BHM, therefore their distributions does not need to be exactly alike. In addition, the missing values in the continuous BHM case were marginalised, while in the discrete BHM were imputed.

On the other hand, the differences between the discrete distributions, BHM and that of Bouy et al. (2015b), arise mainly at the bright and faint ends ($K_s \approx 4$ mag and $K_s \approx 11$ mag). We argue that the origin of these differences lay in our new candidate members and in the rejected ones of Bouy et al. (2015b), as it is discussed in Sect. **??**.

Figure 2.12: Luminosity functions from J, H, K_s (orange). Also shown the regions of incompleteness and the luminosity functions computed from: the candidate members of Bouy et al. (2015b) (dot-dashed blue line), and our candidate members, ($p_{84\%} > p_t$, dashed black line).

.

2.6   Mass distribution

This Section starts with a brief description of the mass-luminosity relation, which is used to derive the present day system mass function (PDSMF) of the Pleiades together with the luminosity distribution of the previous section. Then I compare this PDSMF to the Initial Mass Functions (IMFs) of Chabrier (2005) and Thies & Kroupa (2007). Later, I conclude this section with an analysis of some simple models that can be fitted to our the PDSMF, I give the best model according to the Bayesian evidence.

2.6.1   The mass-luminosity relation

The mass-luminosity relation is the key to transform luminosities into masses. This relation relies entirely on the current models of stellar atmosphere and evolution. Among the different flavours of models, we choose the BT-Settl model of Allard et al. (2012). We based this decision on the fact that currently, this is the only model which covers our luminosity range at the age of the Pleiades. Since the DANCe survey uses as reference the J, H and K_s bands of the 2MASS survey, I choose the BT-Settle grid: CIFIST2011bc for the 2MASS AB photometric system. This grid, as it name indicates, returns values of the luminosity on a (non-uniform) grid on the mass. However, the transformation of luminosities into masses is proportional and thus very sensitive to the Jacobian of the transformation. To avoid the discontinuities in the derivatives produced by the grid, we decided to fit the grid using spline series and then obtain the derivative from these series. It is important to notice that we implicitly assume that the transformation from luminosities to masses does not have any associated uncertainty, thus it is uniquely determined. We can not do more since the models do not provide any uncertainty.

2.6.2   Present day mass function

The PDSMF is independently obtained in the J, H, K_s bands by transforming the luminosity functions into system mass functions using the mass-luminosity relations described in the previous section. Since the luminosity functions of Sect. **??** correspond to the luminosity of systems (single and binary stars altogether), therefore the derived mass function corresponds to the system mass function. We assume an age of 120 Myr for the Pleiades together with solar metallicity, these are the canonical values reported by and , respectively.

Figure **??** shows the logarithmic PDSMF (ξ_L) for the J, H, K_s bands normalised on the completeness limits obtained in Sect.   **??**. This figure also shows, the PDSMF proposed by Bouy et al. (2015b) and, the IMFs of Thies & Kroupa (2007) and Chabrier (2005). For this last IMF, I show the standard uncertainty as the value reported by Chabrier (2003). As shown in this Fig., our PDSMFs compares well, in the completeness interval, with the one proposed by Bouy et al. (2015b). However, there are discrepancies, particularly above $0.3 M_\odot (-0.5 < \log M/M_\odot)$. These may have its origin on the higher membership probability of our new candidate members. These new candidate members are preferentially M stars, whose masses are the range $0.075 - 0.6 M_\odot$ $(-1.12 < \log M/M_\odot < -0.22)$. Furthermore, these differences root also on the objects that Bouy et al. (2015b) did not include in their analysis: those lying outside the inner three degree region.

To obtain a model for our PDSMF I proceed as follows. First, I select three models: a log-normal distribution and two power-law distributions of the form $m^{-\alpha}$, with two and three power-law segments. Second, I took the mode distribution of our 100 sample distributions of the PDSMF in the K_s band and completeness interval and draw a 10^4 synthetic sample. Afterwards, using *PyMultiNest* (Buchner et al. 2014), and the synthetic sample I fitted the three models. Finally, in Table **??**, I give the posterior distributions of each model parameters together with its evidence (see Sect. **??**).

Judging by the evidences, the best model corresponds to the two segment power-law distribution. This model is over plot as solid black line in Fig. **??**. This model agrees well with that found by Bouy et al. (2015b), except for the flat part in the low-mass range and the less step slope in the high mass range. Nevertheless, it is in clear discrepancy with the IMFs of Chabrier (2005), ($m_c = 0.25^{+0.021}_{-0.016}$ and

$\sigma = 0.55^{+0.05}_{-0.01}$, the uncertainties are those reported by Chabrier 2003, for single objects) and of Thies & Kroupa (2007). The discrepancy between the IMFs and the PDSMFs (Bouy et al. 2015b, and ours) may have its origin on the not yet established uncertainties in the mass-luminosity relationship, on dynamical effects associated with age, or on both of them. In the next section I compare the PDSMF of the Pleiades with that of other younger and older clusters in order to analyse if there is substantial evidence to claim for dynamical effects associated with age.

Table 2.1: Parameters and evidence of models fitted to the PDSMF

Model	Parameters		Log Evidence
LogNormal	$m_c = 0.36 \pm 0.03$		
	$\sigma = 0.46 \pm 0.02$		18.1 ± 0.1
Two Segments	$\alpha_0 = -0.11 \pm 0.06$	$m \in [0.04, 0.22 \pm 0.01]$	
	$\alpha_1 = 1.13 \pm 0.1$	$m \in [0.22 \pm 0.01, 0.56]$	2222.7 ± 0.4
Three Segments	$\alpha_0 = -0.05 \pm 0.6$	$m \in [0.04, 0.08 \pm 0.03]$	
	$\alpha_1 = -0.1 \pm 0.1$	$m \in [0.08 \pm 0.03, 0.22 \pm 0.01]$	
	$\alpha_2 = 1.13 \pm 0.1$	$m \in [0.22 \pm 0.01, 0.56]$	2221.2 ± 0.3

Before concluding this section, I use our PDSMF to give a lower limit to the mass of the cluster. The cluster members mean mass, in our entire mass range, is $0.26 \pm 0.006 M_\odot$. Thus, the product of this mean mass with the expected number of cluster members, 3116 ± 110, is an estimate of mass of the cluster: $807^{+38}_{-29} M_\odot$. The expected number of cluster members is the integral, over the whole range of membership probabilities, of number of objects at each membership probability. However, since we still lack the low and high mass ranges of the PDSMF, this value is only a lower limit to the mass of the cluster. As mentioned in Sect. **??**, the uncertainties in the mass-luminosity relations are yet to be established. Thus the quoted uncertainties of our mass results are underestimated.

Figure 2.13: Normalised PDSMF in J, H, K_s bands. Also shown the IMFs of Chabrier (2005); Thies & Kroupa (2007) and fits to the PDSMF found by us and Bouy et al. (2015b).

.

2.7   The mass distribution on time

In order to test if dynamical effects could be the origin of the discrepancies between the PDSMFs and the IMFs, in this section I analyse the evolution in time of the mass distribution. To do this, I compare the differences between the Pleiades PDSMF (≈ 120 Myr) and those of the Trapezium and Hyades clusters, which are ≈ 1 Myr and ≈ 600 Myr old respectively. These can be thought as snapshots of the Pleiades pasts and future mass functions.

In Figure **??** I compare the PDSMF from the Pleiades derived here, to those of the Trapezium and Hyades. These PDSMFs correspond to those of Fig. 11 of Bouy et al. (2015b) (private communication). As mentioned by Bouy et al. (2015b), the abundance of low-mass stars and brown dwarfs in the range $0.03 - 0.1$ $M_\odot (\log M/M_\odot \approx \{-1, -1.4\})$ seems to diminish with time. Furthermore, since the PDSMF are normalised, a diminish in the lower mass range produces a relative increase of low-mass stars in the range $-0.4 < \log M/M_\odot < -0.2$. The alternative effect of an increase in the high mass range which then may produce a relative diminish in the low mass range, although is statistically possible, is unlikely in the astrophysical sense. Open clusters lose stars, not the other way around. Furthermore, this effect is consistent with the classical scenario in which low-mass stars and brown dwarfs are ejected as the cluster relaxes. I test the validity of this scenario, at least the statistical significance of the observed differences among the PDSMF of this three clusters. To perform this test, the null hypothesis is that the Trapezium and the Hyades have the same PDSMF as the Pleiades. Since I only have the inferred model of the Pleiades cluster, I am compelled to perform a frequentist test. Thus, to do the statistical comparison of these three PDSMF I use the Kolmogorov-Smirnov and the Anderson-Darling tests.

Figure 2.14: Left: PDSMFs of the Pleiades (derived here for J, H, K_s bands), Trapezium, and Hyades, from Bouy et al. (2015b). They are normalised in the interval of completeness.

.

Figure **??** shows the cumulative distribution functions (CDFs) of the Trapezium, Pleiades (only in K_s band) and Hyades PDSMFs. Also and for comparison,

Figure 2.15: Cumulative distribution functions (CDF) of the PDSMFs from left panel and that of Chabrier (2005) and Thies & Kroupa (2007) system initial mass function (normalised also in the interval of completeness). The Pleiades CDF shown is just from K_s band. The grey area depicts the area in which the null hypothesis of same PDSMF as that of the Pleiades can not be rejected (at $\alpha = 0.01$).

.

we show the CDFs of Chabrier (2005) and Thies & Kroupa (2007) IMFs. The grey area around the Pleiades CDF shows the hypothesis test in which I compare each CDF with that of the Pleiades. The null hypothesis,as mentioned before is that each compared CDF is exactly that of the Pleiades. I use the Kolmogorov-Smirnov statistic and the alpha value $\alpha = 0.01$, to compute the maximum vertical distance d_α from the Pleiades CDF, the grey region was created with this maximum distance. The null hypothesis is rejected only if the tested CDF lies entirely outside the grey region around the Pleiades CDF. As can be seen, neither the IMFs nor the PDSMF of the Trapezium and Hyades lay entirely within the grey area, thus we can reject the null hypothesis that they share the same PDSMF of the Pleiades. Furthermore, since the Kolmogorov-Smirnov test uses only the maximum distance between CDFs, we also applied the more robust Anderson-Darling test. It also rejects the null hypotheses (at $p < 0.004$) that the Trapezium and Hyades PDSMFs and the Chabrier (2005) and Thies & Kroupa (2007) IMFs have the same CDF of the Pleiades.

The previous tests show that there is enough evidence to claim for differences among the PDSMFs of these three clusters and from IMFs and Pleiades PDSMF. Thus suggesting that these differences may have an origin on dynamical effects associated with age and relaxation. Nevertheless, to claim for reliable evidence supporting these differences the census of the Trapezium and Hyades must be done using the same methods. Also, the uncertainties must be properly establish both for the other PDSMFs and for the mass-luminosity relation from which all these PDSMF are derived.

Bibliography

????

Adams, J. D., Stauffer, J. R., Monet, D. G., Skrutskie, M. F., & Beichman, C. A. 2001, AJ, 121, 2053

Akeret, J., Seehars, S., Amara, A., Refregier, A., & Csillaghy, A. 2013, Astronomy and Computing, 2, 27

Allard, F., Homeier, D., Freytag, B., & Sharp, C. M. 2012, in EAS Publications Series, Vol. 57, EAS Publications Series, ed. C. Reylé, C. Charbonnel, & M. Schultheis, 3–43

Alsing, J., Heavens, A., & Jaffe, A. H. 2017, MNRAS, 466, 3272

Anderes, E., Wandelt, B. D., & Lavaux, G. 2015, ApJ, 808, 152

Andrieu, C., de Freitas, N., Doucet, A., & Jordan, M. I. 2003, Machine Learning, 50, 5

Azzalini, A. 1996, Biometrika, 83, 715

Azzalini, A. & Capitanio, A. 1999, Journal of the Royal Statistical Society: Series B (Methodological)

Balaguer-Núñez, L., Galadí-Enríquez, D., & Jordi, C. 2007, A&A, 470, 585

Barnes, III, T. G., Moffett, T. J., Jefferys, W. H., & Forestell, A. D. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 310, IAU Colloq. 193: Variable Stars in the Local Group, ed. D. W. Kurtz & K. R. Pollard, 95

Barrado, D., Bouy, H., Bouvier, J., et al. 2016, A&A, 596, A113

Bishop, C. 2006, Pattern Recognition and Machine Learning, Information Science and Statistics (Springer)

Blackwell, T. & Bentley, P. 2002, in Proceedeings of the 4th Annual Conference on Genetic and Evolutionary Computations, ed. W. Langdon, E. Cantu-Paz, K. Mathias, R. Roy, & D. Davis (San Francisco, CA, USA: Morgan kaufmann Publishers Inc.), 19–26

Bouvier, J., Kendall, T., Meeus, G., et al. 2008, A&A, 481, 661

Bouy, H., Bertin, E., Barrado, D., et al. 2015a, A&A, 575, A120

Bouy, H., Bertin, E., Moraux, E., et al. 2013, A&A, 554, A101

Bouy, H., Bertin, E., Sarro, L. M., et al. 2015b, A&A, 577, A148

Bovy, J., Hogg, D. W., & Roweis, S. T. 2011, The Annals of Applied Statistics, 5, 1657

Brooks, S., Gelman, A., Jones, G., & Meng, X. 2011, Handbook of Markov Chain Monte Carlo, Chapman & Hall/CRC Handbooks of Modern Statistical Methods (CRC Press)

Buchner, J., Georgakakis, A., Nandra, K., et al. 2014, A&A, 564, A125

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245

Chabrier, G. 2003, PASP, 115, 763

Chabrier, G. 2005, in Astrophysics and Space Science Library, Vol. 327, The Initial Mass Function 50 Years Later, ed. E. Corbelli, F. Palla, & H. Zinnecker, 41

Chung, Y., Gelman, A., Rabe-Hesketh, S., Liu, J., & Dorie, V. 2015, Journal of Educational and Behavioral Statistics, 40, 136

Clerc, M. & Kennedy, J. 2002, Trans. Evol. Comp, 6, 58

Converse, J. M. & Stahler, S. W. 2008, ApJ, 678, 431

D'Antona, F. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 142, The Stellar Initial Mass Function (38th Herstmonceux Conference), ed. G. Gilmore & D. Howell, 157

De Boor, C. 1978, A practical guide to splines / Carl de Boor (Springer-Verlag New York), xxiv, 392 p. :

Demory, B.-O. 2014, ApJ, 789, L20

Dempster, A. P., Laird, N. M., & Rubin, D. B. 1977, JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B, 39, 1

Diard, J. & Bessiere, P. 2008, Probabilistic reasoning and decision making in sensory-motor systems, 153

Dirac, P. A. M. 1942, Proceedings of the Royal Society of London Series A, 180, 1

Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. 1987, Physics Letters B, 195, 216

Duquennoy, A. & Mayor, M. 1991, A&A, 248, 485

Faherty, J. K., Burgasser, A. J., Walter, F. M., et al. 2012, ApJ, 752, 56

Fan, Y. & Sisson, S. 2011, Handbook of Markov Chain Monte Carlo, 67

Feeney, S. M., Johnson, M. C., McEwen, J. D., Mortlock, D. J., & Peiris, H. V. 2013, Phys. Rev. D, 88, 043012

Fei-Fei, L. & Perona, P. 2005, in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Vol. 2, Ieee, 524–531

Feroz, F., Hobson, M. P., & Bridges, M. 2009, Monthly Notices of the Royal Astronomical Society, 398, 1601

Flegal, J. M., Hughes, J., & Vats, D. 2016, mcmcse: Monte Carlo Standard Errors for MCMC, Riverside, CA and Minneapolis, MN, r package version 1.2-1

for Guides in Metrology, J. C. 2008, JCGM 100: Evaluation of Measurement Data - Guide to the Expression of Uncertainty in Measurement, Tech. rep., JCGM

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306

Fraley, C. & Raftery, A. E. 2002, Journal of the American Statistical Association, 97, 611

Fraley, C., Raftery, A. E., Murphy, T. B., & Scrucca, L. 2012, mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation

Franklin, J. 2002, The Science of Conjecture: Evidence and Probability Before Pascal, A Johns Hopkins paperback (Johns Hopkins University Press)

Gagné, J., Lafrenière, D., Doyon, R., Malo, L., & Artigau, É. 2014, ApJ, 783, 121

Galli, P. A. B., Moraux, E., Bouy, H., et al. 2017, A&A, 598, A48

Gelman, A. 2006, Bayesian Analysis, 1, 515

Gelman, A. 2012, CHANCE, 25, 52

Gelman, A., Carlin, J., Stern, H., et al. 2013, Bayesian Data Analysis, third edition edn. (CRC)

Gelman, A. & Hill, J. 2007, Data Analysis Using Regression and Multilevel/Hierarchical Models, Analytical Methods for Social Research (Cambridge University Press)

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. 2008, Ann. Appl. Stat., 2, 1360

Geman, S. & Geman, D. 1984, IEEE Trans. Pattern Anal. Mach. Intell., 6, 721

Gong, L. & Flegal, J. M. 2016, Journal of Computational and Graphical Statistics, 25, 684

González-Farias, G., Dominguez-Molina, A., & Gupta, A. K. 2004, Journal of Statistical Planning and Inference, 126, 521

Good, I. J. 1980, Trabajos de Estadistica Y de Investigacion Operativa, 31, 489

Goodman, J. & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Gupta, A. K., González-Farias, G., & Dominguez-Molina, J. 2004, Journal of Multivariate Analysis, 89, 181

Guthrie, B. N. G. 1987, QJRAS, 28, 289

Hand, D. J. & Yu, K. 2001, International Statistical Review / Revue Internationale de Statistique, 69, 385

Hastings, W. K. 1970, Biometrika, 57, 97

Hogg, D. W., Bovy, J., & Lang, D. 2010a, ArXiv e-prints

Hogg, D. W., Myers, A. D., & Bovy, J. 2010b, ApJ, 725, 2166

Hong, D., Balzano, L., & Fessler, J. A. 2016, Towards a theoretical analysis of PCA for heteroscedastic data, arxiv 1610.03595

Huang, A. & Wand, M. P. 2013, Bayesian Analysis, 8, 439

Jaynes, E. 2003, Probability Theory: The Logic of Science (Cambridge University Press)

Jefferys, T. R., Jefferys, W. H., Barnes, III, T. G., & Dambis, A. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 371, Statistical Challenges in Modern Astronomy IV, ed. G. J. Babu & E. D. Feigelson, 433

Kass, R. E. & Raftery, A. E. 1995, Journal of the American Statistical Association, 90, 773

Kennedy, J. & Eberhart, R. 1995, in Neural Networks, 1995. Proceedings., IEEE International Conference on, Vol. 4, 1942–1948 vol.4

Koller, D. & Friedman, N. 2009, Probabilistic Graphical Models: Principles and Techniques, Adaptive computation and machine learning (MIT Press)

Krone-Martins, A. & Moitinho, A. 2014, A&A, 561, A57

Lin, T., Lee, J., & Ho, H. 2006, WOS:000236696400015

Lindstrom, M. 1999, Journal of Computational and Graphical Statistics, 8, 333

Liu, Z., Chen, Y., Tian, S., & Xu, Z. 2015, Journal of Information and Computational Science, 12, 775

Lodieu, N., Deacon, N. R., & Hambly, N. C. 2012, MNRAS, 422, 1495

Malo, L., Doyon, R., Lafrenière, D., et al. 2013, ApJ, 762, 88

March, M. C., Karpenka, N. V., Feroz, F., & Hobson, M. P. 2014, MNRAS, 437, 3298

McMichael, D. W. 1996, in in Proc. Fourth International Symposium on Signal Processing and its Applications (ISSPA, 377–378

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, The Journal of Chemical Physics, 21, 1087

Moraux, E., Bouvier, J., & Stauffer, J. R. 2001, A&A, 367, 211

Morris, C. N. 1983, Journal of the American Statistical Association, 78, 47

Muench, A. A., Lada, E. A., Lada, C. J., & Alves, J. 2002, ApJ, 573, 366

Neal, R. M. 1996, Bayesian Learning for Neural Networks (Secaucus, NJ, USA: Springer-Verlag New York, Inc.)

Olivares, J. & et al. 2017a, submitted to å

Olivares, J. & et al. 2017b, in preparation

Ozerov, A., Lagrange, M., & Vincent, E. 2013, Computer Speech and Language, 27, 874 , special Issue on Speech Separation and Recognition in Multisource Environments

Patel, P. K., Sharma, V., & Gupta, K. 2013, International Journal of Computer Applications, 73

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825

Press, W. H. 1997, in Unsolved Problems in Astrophysics, ed. J. N. Bahcall & J. P. Ostriker, 49–60

Pugachev, V. 1965, Theory of Random Functions and its Application to Control Problems (Pergamon), 852

Riedel, A. R., Blunt, S. C., Lambrides, E. L., et al. 2017, AJ, 153, 95

Roberts, G. O. & Rosenthal, J. S. 2004, Probability Surveys, 1, 20

Sale, S. E. 2012, MNRAS, 427, 2119

Sampedro, L. & Alfaro, E. J. 2016, MNRAS, 457, 3949

Sarro, L. M., Bouy, H., Berihuete, A., et al. 2014, Astronomy & Astrophysics, 14

Schwarz, G. 1978, Ann. Statist., 6, 461

Shariff, H., Dhawan, S., Jiao, X., et al. 2016, MNRAS, 463, 4311

Shkedy, Z., Decin, L., Molenberghs, G., & Aerts, C. 2007, MNRAS, 377, 120

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, American Institute of Physics Conference Series, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405

Skilling, J. 2006, Bayesian Anal., 1, 833

Smyth, G. K. 1998, in Encyclopedia of Biostatistics, ed. Armitage, P. & Colton, T. (John Wiley and Sons)

Spiriti, S., Eubank, R., Smith, P. W., & Young, D. 2013, Journal of Statistical Computation and Simulation, 83, 1020

Stauffer, J. R., Hartmann, L. W., Fazio, G. G., et al. 2007, ApJS, 172, 663

Stauffer, J. R., Schultz, G., & Kirkpatrick, J. D. 1998, ApJ, 499, L199

Stoneking, C. J. 2014

Takeda, Y., Hashimoto, O., & Honda, S. 2017, PASJ, 69, 1

Tapiador, D., Berihuete, A., Sarro, L. M., Julbe, F., & Huedo, E. 2017, Astronomy and Computing, 19, 1

Thies, I. & Kroupa, P. 2007, ApJ, 671, 767

Trumpler, R. J. 1921, Lick Observatory Bulletin, 10, 110

Wolfgang, A. & Lopez, E. 2015, ApJ, 806, 183